DOCUMENT RESUME

ED 100 991 TM 004 112

AUTHOR	Linn, Bernard S.; And Others
TITLE	Peer and Self Assessment in the Quest for Evaluative Techniques That Predict Delivery of Quality Care.
SPONS AGENCY	Veterans Administration, Washington, D.C.
PUB DATE	[Nov 73]
NOTE	7p.; Paper presented at the Annual Conference on Research in Medical Education (Washington, D. C., November 1973)
EDRS PRICE	MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS	Evaluation Methods; *Medical Students; *Medical Treatment; *Peer Groups; *Performance; Prediction; *Self Evaluation
IDENTIFIERS	Physician Performance Rating Scale

ABSTRACT

Ninety-eight medical students rated themselves and their fellow students using the Physician Performance Rating Scale (PPRS). Two factors were studied: a knowledge factor (judgement, ability) and a relationship factor (appearance, interpersonal relationships, ability to communicate). Reliability between peer and self ratings was significant but low. Students rated themselves generally lower than did their peers, especially in interpersonal skills. Peer ratings were highly significantly related to grade actually attained. The relationship factor was inversely related to grade outcome. (SM)

ED100691

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PEER AND SELF ASSESSMENT IN THE QUEST FOR EVALUATIVE TECHNIQUES THAT PREDICT DELIVERY OF QUALITY CARE[1]

BERNARD S. LINN, M.D.
Associate Professor of Surgery
University of Miami School of Medicine
and Veterans Administration Hospital
Miami, Florida

MARTIN AROSTEGUI, M.D.
Medical Resident
University of Miami School of Medicine
Miami, Florida

ROBERT ZEPPA, M.D.
Professor and Chairman
Department of Surgery
University of Miami School of Medicine
Miami, Florida

TM 004 112

[1]Paper presented at the 12th Annual Conference on Research in Medical Education, November 8, 1973, Washington, D.C.

Work supported in part by VA 8200 Research and Education funds.

The ultimate criterion of a physician's performance is the quality of care which he delivers. Much has been written about measurement of these two variables. Even if quality of care can be estimated accurately, the intervening years between medical school and practice make it difficult to identify markers of success. Grades reportedly (1) have had little predictive validity, neither has our system of credentialing (2,3). Although student behaviors seem important, those which specifically predict success in practice have not been clearly identified. Furthermore, there is still a question concerning how behaviors can best be measured and by whom.

With the PSROs, recertification, and documentation of continued competence, the focus for evaluation has moved beyond the medical student to include the resident and the physician throughout his clinical practice. Some emphasis has been given by the AMA recently to self assessment in continuing education. Little attention has been given to this type of evaluation in medical school. Yet, it would seem that helping the student learn to evaluate himself would be an important asset to him throughout his career. With the interest in peer review in judging quality of care, it is also surprising that peer evaluation has not been used more in measuring student performance but where it has (4-7), it has been reported of value.

The purpose of this study was to determine if students could identify their own strengths and weaknesses as viewed by their peers and as reflected in other outcome measures such as grades and National Board Examinations (NBE).

METHOD

Ninety-eight of 102 junior medical students participated in the study. Ratings were taken on a Physician Performance Rating Scale (8) containing 16 items for peer and 13 items for self evaluation. Items were scored on 4-point scales from below average to outstanding, with higher scores being more favorable. The scale yields two factors: A Knowledge Factor of 6 items (fund of knowledge, clinical judgment, intellectual curiosity, technical ability, potential as a physician, and desirability as a

as a consultant), and a Relationship Factor with 10 items (conscientiousness, leadership, professional appearance, doctor-patient relationship, team spirit, interpersonal relationships, integrity, ability to communicate ideas, desirability as a partner, and desirability as one's own physician). The Knowledge Factor has a test-retest value of r=.907 for peers and .680 for self. The Relationship factor has an r=.902 for peers and .824 for self.

Students rotated through their surgical clerkship in small groups. Each student rated himself and his peers at the end of the 12 week rotation. This provided 928 ratings. Ratings from all peers were averaged for each student to compare with his own evaluation of himself on the two factor scores.

Average grade on quiz and essay exams, grade given by house staff, grade by tutor, final grade in surgery, and scores on the surgical section of NBE were used as other outcome variables. Ratings from peers were compared with self ratings by Pearson correlation. Peer and self ratings were used to predict all other grade type outcomes singly and in combination, by multivariate analysis of covariance, and all measures were factor analysed to determine the dimensions in evaluation.

RESULTS

Reliability between peer and self ratings was significant but low. On Knowledge, peer and self correlated at $P<.01$ ($r=.408$). For Relationship, the correlations was only $r=.297$ which was statistically significant at the 5% level. Self ratings tended to be lower than those assigned to the person by his peers, indicating students rated themselves as less favorable. In addition, students and peers agreed less about their interpersonal skills than they did about their knowledge.

Table 1 shows how peer and self factors predicted all grade type variables taken in combination. Only peer ratings were significantly related to these variables and at highly significant levels. The Relationship Factor was related to grade outcome inversely. In other words, those judged by peers to perform best in interpersonal skill received significantly lower overall evaluation.

BEST COPY AVAILABLE

The second part of Table 1 shows self and peer factors as they predicted scores on the surgical section of NBE, alone. Both peer and self Knowledge Factors predicted scores on NBE significantly. Although the Relationship Factors are not statistically significant, they are again inversely related to outcome.

The last part of Table 1 demonstrates how these same variables related to final grade in surgery. Final grade was based on results from quizzes, essays, NBE, and ratings from faculty tutors and house staff. All factors were significant predictors

(INSERT TABLE 1 ABOUT HERE)

of final grade, with peer ratings being the most predictive at a .001 level and self ratings at only $P < .05$. Again, both Relationship Factors are inversely correlated with final grade.

Lastly, Table 2 gives the factor structure of the variables used in evaluation. Four rather strong factors emerged, suggesting independent dimensions within the vari-

(INSERT TABLE 2 ABOUT HERE)

ables. Essays, NBE, house staff grades, and final grade form the first factor accounting for 25% of the variance, peer ratings for 21%, self ratings for 18%, and quiz and faculty for only 10%.

DISCUSSION

Firstly, students were more critical of themselves than their peers were of them. A case might be made for that being desirable; however, peer assessments were always better predictors of final grade, NBE, and the multivariate combination of all grade-type outcomes. This could mean either that peer ratings were more valid or just that peers tended to judge students similarly to the way they were judged by faculty. Secondly, behavior tended to relate inversely to grades. Geertsma and Chapman (9) reported similarly finding two factors (cognitive and non-cognitive) and showed students who scored high on one scored low on the other. This is somewhat disturbing. Does it mean that students with good behavior are less intelligent? Or, does it imply that popular students are busy socializing and into extra-curricular activities and

BEST COPY AVAILABLE

thus spend less time studying?

It would seem from this study that students are able to judge knowledge of their peers, in fact, even better than faculty. There is no reason to believe they do not make equally as adequate an assessment about personality related behaviors and skills. The fact that interpersonal behaviors predicted grades inversely could mean that medicine has worshiped too long at the alter of academic achievement and therefore has passed the counds of what may be needed and necessary in a practicing physician. Most students entering medical school are already in the upper IQ levels, increasingly so each year, and the tendency toward "elitism" in scholastic achievement may not be related to high level performance as a clinician, but more predictive of careers in medical research where interpersonal skills are not as vital. It would seem that the major emphasis should be toward defining and measuring the quality of medical care, for if this end point of success is measurable, we may then be able to determine accurate predictors.

REFERENCES

1. WINGARD, J.R. and WILLIAMSON, J.W. Grades as Predictors of Physicians' Career Performance: An Evaluative Literature Review. J. Med. Ed. 48:311-322, 1973.

2. PETERSON, O.L.,ANDREWS, L.P., SPAIN, R.S., and GREENBERG, B.G. An Analytical Study of North Carolina General Practice, J. Med. Ed. 31:1-165, 1956.

3. PRICE, P.B., TAYLOR, C.W., RICHARDS, J.M., and JACOBSEN, T.C. Measurement of Physician Performance. J. Med. Ed. 39:203-211, 1964.

4. KORMAN, M., STUBBLEFIELD, R.L., and MARTIN, L.W. Patterns of Success in Medical School. J. Med. Ed. 43:405-411, 1968.

5. SCHUMACHER, C.F. A Factor-Analytic Study of Various Criteria of Medical Examinations. J. Med. Ed. 39:192-196, 1964.

6. HAMMOND, K.R. and KERN, F. Teaching Comprehensive Medical Care, Harvard Univ. Press, Cambridge, 1959.

7. HELFER, R.E. Peer Evaluation: Its Potential Usefulness in Medical Education. Br. J. Med. Ed. 6:224-231, 1972.

8. LINN, B.S., AROSTEGUI, M., and ZEPPA, R. A Performance Rating Scale for Peer and Self Assessment. (in press)

TABLE 1

PREDICTION OF SURGICAL CLERKSHIP OUTCOME BY PEER AND SELF FACTOR SCORES

VARIABLES	B WEIGHT	F-RATIO	P LEVEL
GRADE TYPE VARIABLE TOGETHER			
Self Relationship	-.212	.497	.776
Self Knowledge	.132	1.100	.378
Peer Relationship	-.372	4.167	.003***
Peer Knowledge	1.272	8.828	.001***
SURGICAL SECTION OF NBE			
Self Relationship	-.581	.606	.439
Self Knowledge	.431	3.961	.051
Peer Relationship	-1.904	1.026	.315
Peer Knowledge	3.532	15.489	.001***
FINAL GRADE IN SURGERY			
Self Relationship	-.260	4.113	.047*
Self Knowledge	.030	6.245	.015*
Peer Relationship	-.653	8.370	.001***
Peer Knowledge	2.495	19.978	.001***

*P < .05
**P < .01
***P < .001

BEST COPY AVAILABLE

TABLE 2

FACTOR STRUCTURE OF ALL EVALUATION MEASURES

VARIABLES	FACTOR I	FACTOR II	FACTOR III	FACTOR IV
Quiz				.698
Essay	.744			
NBE	.670			
Resident Eval.	.542			
Faculty Eval.				-.664
Final Grade	.932			
Self Relationship		.879		
Self Knowledge		.896		
Peer Relationship			.913	
Peer Knowledge			.855	